Radio signals of cosmic-ray-induced air showers at the Pierre Auger Observatory

Harmsma, Sybren

6 Observations

In this chapter, some general properties of the data will be discussed. First, an overview will be given of the air showers that were selected from the Auger Surface Detector (SD). Next, general aspects of the data that was measured with the first-generation radio setup will be discussed. Topics include calibration, timing precision, self-interference, and the radio background at the site. The radio measurements of actual cosmic-ray induced air showers will be the topic of Chapter 7.

6.1 Properties of SD events

6.1.1 Core positions and shower energies

In Section 5.2, “Events selected from the surface detector”, a set of criteria was defined for the events from the Auger SD. Events that fulfilled all the criteria were selected for analysis. This selection consists of 2,028 air-shower events, and in this section a brief overview of the properties of these events will be given.

One of the criteria by which the SD events were selected, is that the SD station that is nearest to the radio setup, namely Olaia, must be part of the (reconstructed) event. Due to this criterion nearby events are selected. The core positions of these events are shown in Figure 6.1. In this thesis, the terms 'core position' and 'shower core' indicate the location where the shower axis reaches the surface of the Earth.

The peculiar shape of the event distribution in this figure may seem surprising at first. However, it can be understood by comparing it with the positions of the SD stations. The triangular grid of the SD stations is also plotted in Figure 6.1. The off-grid position of the “Olaia” SD station lowers the energy threshold of detectable air showers since it allows smaller showers to be detected. Since the cosmic-ray energy spectrum is very steep, as was shown in Figure 1.4 on p. 7, this results in a large number of additional, low-energy events.

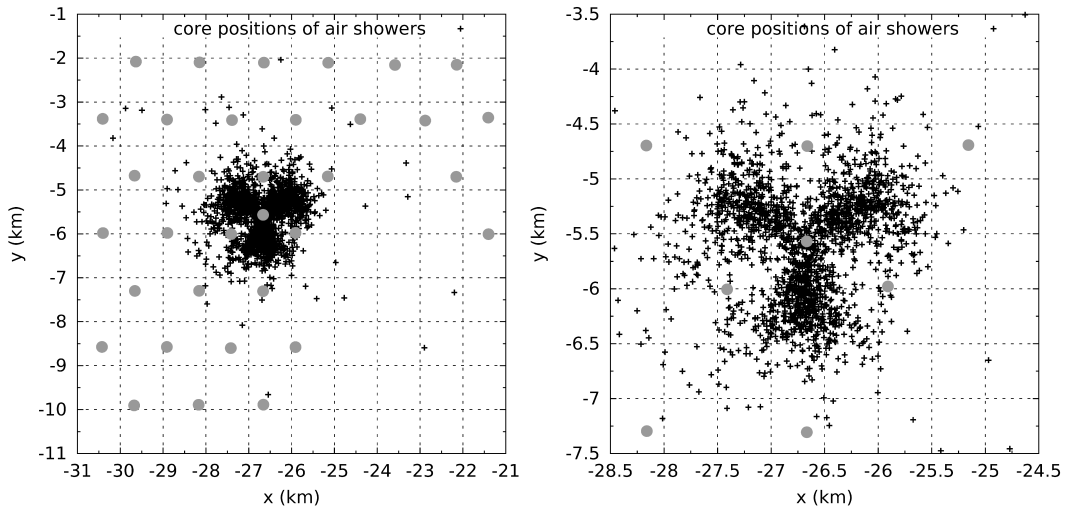

Figure 6.1: The x- and y-coordinates of the shower cores of selected SD events. The x-coordinate is positive towards the east; y is positive towards the north. The origin of the coordinate system is located at the centre of the Auger observatory. The positions of nearby SD tanks (detectors) have been indicated with grey circles. Left: a display of all 2028 selected events. Right: a zoom-in of the central region.

Figure 6.2 shows that the majority of the selected SD events (73 %) has an energy[2] of less than 0.5 EeV. Further inspection showed that 86 % of these low-energy events were detected with only 3 SD stations. The smallest configurations that consist of 3 SD stations are indicated with ellipses in Figure 6.3 (a). The centre of each ellipse is indicated with a ⊗ symbol; around these locations we expect a larger number of core positions due to the lower energy threshold.

In the (b), (c) and (d) sub-plots of Figure 6.3, a cut has been made on the energy of the events. The energy is required to be more than 0.3, 0.5 and 1.0 EeV, respectively. The characteristic pattern of Figure 6.1 is no longer visible for events with energies above 0.5 EeV. Thus, the combination of the larger flux of low-energy cosmic rays with the detector positions is the cause of this pattern.

[2]In this thesis, the value we use for the energy of the SD events is the energy that is calculated during the reconstruction with the CDAS software [99]. Specifically, the "NEWFD" energy estimator is used, which uses the constant intensity cut (CIC) method, together with the latest FD calibration [100, 101]. This "NEWFD" energy estimator was at the time of writing the preferred energy estimator in the CDAS reconstruction software.

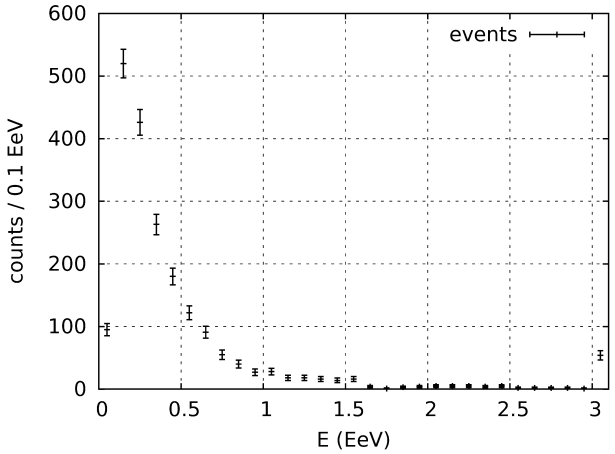

Figure 6.2: Energy distribution of the selected SD events. The bin width is 0.1 EeV (10^{17} eV). The error bars indicate the statistical uncertainty only. The final bin shows 54 events that had a reconstructed energy of more than 3 EeV.

It is necessary, however, to exercise some caution when drawing conclusions based on the SD data. The majority of the events (70 %) has been detected with only 3 SD stations. The uncertainty in the core position and energy of these events is relatively large. Other reconstruction parameters, such as the zenith and azimuth angles of the arrival direction, do also have a larger uncertainty on them as a result of the small number of detectors with which they were measured. Fortunately, the event reconstruction procedure of CDAS provides us with an estimate of the uncertainty in the reconstruction parameters. An overview of the estimated uncertainties of the selected SD events is given in Table 6.1.

No. of SD stations	**No. of events**	$\overline{E}$ (EeV)	$\overline{dE}$ (EeV)	$\overline{dE/E}$	$\overline{dx_{core}}$ (m)	$\overline{dy_{core}}$ (m)	$\overline{d\theta}$ (°)	$\overline{d\varphi}$ (°)
3	1416	0.28	0.12	0.50	97	92	2.4	5.3
4	401	0.62	0.19	0.34	72	76	1.2	2.1
5	119	1.3	0.33	0.28	60	67	0.95	1.5
≥6	92	4.0	1.21	0.27	112	63	0.69	0.64

Table 6.1: Uncertainty of several reconstructed parameters as a function of the number of SD stations that were used in the reconstruction. θ indicates the zenith angle; φ is the azimuth angle.

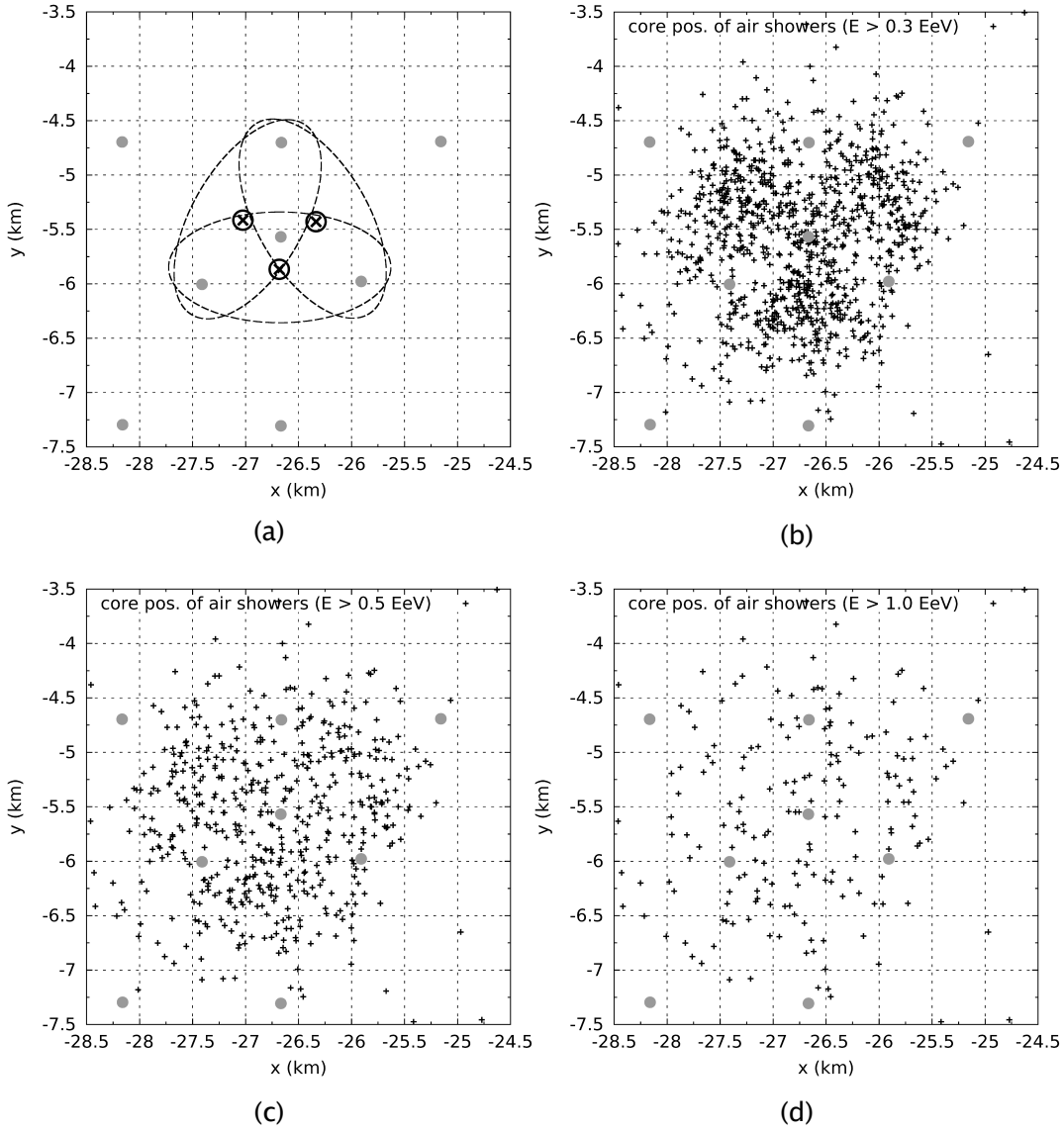

Figure 6.3: Dependency of the shower core position on energy. (a) Sketch of smallest 3-tank configurations. (b) When applying and energy cut E > 0.3 EeV, the characteristic pattern (as seen in Figure 6.1) is still visible. (c) For energies E > 0.5 EeV this pattern is no longer visible. (d) Idem for E > 1.0 EeV.

When comparing high-energy events with low-energy events, we find that high-energy events have a smaller relative uncertainty in their energy. Additionally, high-energy events also have a smaller uncertainty in both the zenith and azimuth angles. The uncertainty in the x- and y-coordinates of the shower core becomes smaller if the number of detectors with which the event was measured, is increased. The average uncertainty in the x-coordinate for events measured with 6 or more SD stations, as listed in Table 6.1, is a notable

exception to this rule. This however is only an artefact that is caused by a few (seven) events with a very large uncertainty in the x-coordinate.

The relative uncertainty in the energy is around 50 % for events that were measured with only 3 SD stations. Naturally, a more closely spaced grid of SD stations would have been able to measure these events with more accuracy. The accuracy of the reconstructed air-shower parameters of few-tank events, such as E and dE, was investigated by Fitzner [102, 103].

The energy distribution of the selected SD events was presented in Figure 6.2. However, it is more common to use a logarithmic scale for the energy axis and to use energy bins that scale exponentially. The resulting flux spectrum is presented in Figure 6.4.

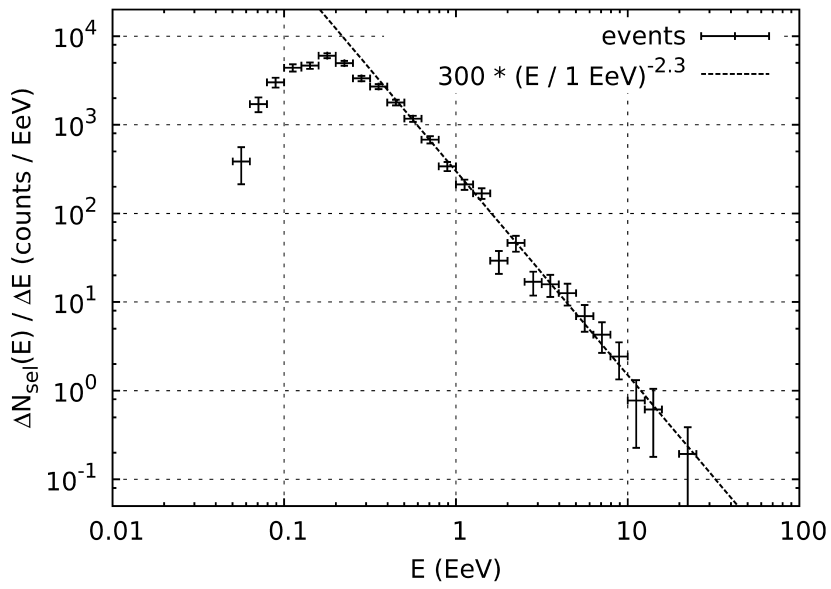

Figure 6.4: Differential flux spectrum of the selected SD events. The vertical error bars indicate the statistical uncertainty only. The top bin, $10^{10.3} \leq E < 10^{10.4}$, contains one event. A power law function is shown to guide the eye. The energy-dependent acceptance has not been taken into account.

In [15], the flux of cosmic rays J as a function of energy E is given by $J(E) \simeq \frac{\Delta N_{sel}(E)}{\Delta E} \frac{1}{\mathscr{E}(E)}$. In this formula, $\Delta N_{sel}(E)$ is the number of selected events in an energy bin centred around E; ΔE is the width of that bin; and $\mathscr{E}(E)$ is the energy-dependent exposure of the detector.

In the referenced article, data from the (full) Auger observatory was used to fit a power law $J = kE^{-\gamma}$ to the cosmic-ray flux spectrum below the ankle (see also Figure 1.4 and Figure 1.5). The value found for γ was $3.28 \pm 0.07(\text{stat})^{+0.11}_{-0.10}(\text{sys})$.

What is plotted in Figure 6.4 is the first part of the formula for $J(E)$: $\Delta N_{sel}(E)/\Delta E$. To guide the eye, a power law has also been plotted in this figure; for this function we used $\gamma = 2.3$. Since the energy-dependent exposure $\mathscr{E}(E)$ is not included in the calculation, we do not expect both values for γ to correspond.

It is difficult, however, to obtain this exposure. In [104] and [105] the acceptance of the regular Auger grid (1500 m triangular; see Section 3.2), is calculated. Both references find that the acceptance saturates (> 99 %) at $E = 2.2\,\text{EeV}$. Additionally, [105] also finds the acceptance for a 750 m triangular grid, which is called “the infill” and is used for AMIGA [106]. It was found that the acceptance of a 750 m grid saturates at $E = 0.15\,\text{EeV}$. However, the SD station “Olaia” is 866 m away from its 3 nearest neighbours, and is not part of an 866 m triangular grid.

Additionally, as can be seen in Figure 6.1 (left), a group of SD stations was missing from the regular grid during the time of the selected SD events. This concerns 12 SD stations, located to the south-east from Olaia. This gap also influences the acceptance. Furthermore, the acceptance is also affected by the criteria which are used for the selection of events. The selection criteria that were used in this thesis for the SD events, as described in Section 5.2, are very different from the criteria which were used in [15]. Finally, as we shall see in Section 6.1.2, the acceptance is also influenced by the atmosphere in the case of low-energy inclined cosmic rays.

6.1.2 Arrival direction

Another feature of the SD events that can be explored is the arrival direction. This arrival direction can be characterized by the zenith and azimuth angles. These angles are defined in the local frame of the Auger observatory, i.e. the rotation of the Earth is not taken into account. The horizontal angle is the azimuth angle φ, which is 0° towards the east (positive x-axis), 90° towards the

north (positive y-axis) and 180° towards the west. The zenith angle θ is 0° for fully vertical events and 90° for fully horizontal events.

The distribution of the zenith angles of the selected SD events has been plotted in Figure 6.5 (left). One apparent feature is that the number of events with a small (< 20°) zenith angle is relatively small. This is not surprising since the amount of solid angle is a function of the zenith angle. The bin $0° < \theta < 2°$ describes a small circle on the sky, while e.g. $40° < \theta < 42°$ covers a much larger ring. To compensate for this dependency, the distribution was calculated as a function of cos(θ); this ensures that each bin covers an equal amount of solid angle. The results are shown in Figure 6.5 (right).

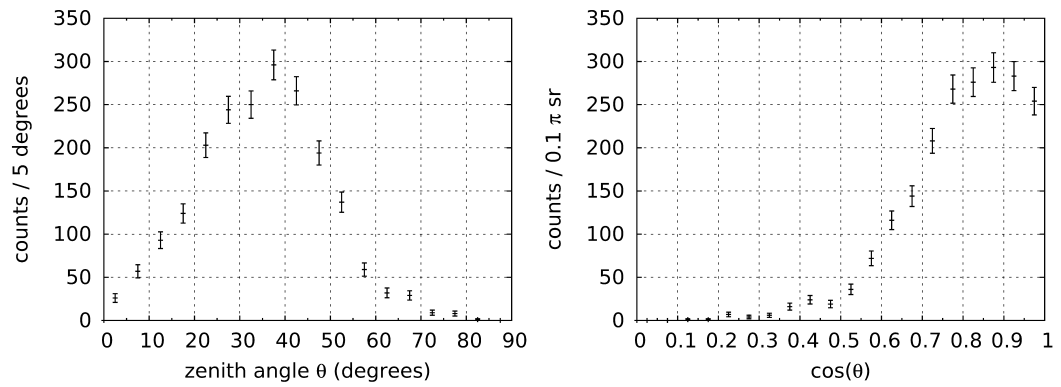

Figure 6.5: Left: distribution of the zenith angles of the selected SD events. The zenith angle is zero in the fully vertical direction. Right: the same distribution, but now as a function of cos(θ). Only statistical uncertainties are shown.

The distribution in Figure 6.5 (right) is fairly constant for zenith angles below about 40° ($\cos(\theta) > 0.75$). Above 40°, the distribution starts to fall off, with almost no events with $\theta > 60°$ ($\cos(\theta) < 0.5$). This reduction can be explained by the larger slant depth in the atmosphere for highly inclined showers (i.e. showers with a large zenith angle). This larger slant depth will cause more air showers to be absorbed before they reach the surface. This principle can be made visible in a scatter plot of the cosmic-ray energy versus the zenith angle; see Figure 6.6.

In this figure many events with low energies (e.g. 0.2 EeV) can be seen at zenith angles smaller than 55°. Above 55°, however, are nearly no low-energy events. The atmospheric absorption appears to be important only for air showers

with a relatively low energy; no suppression can be seen for higher-energy (e.g. E > 3.0 EeV) air showers. (Here we have not considered the case for events that are extremely inclined, i.e. horizontal or nearly horizontal).

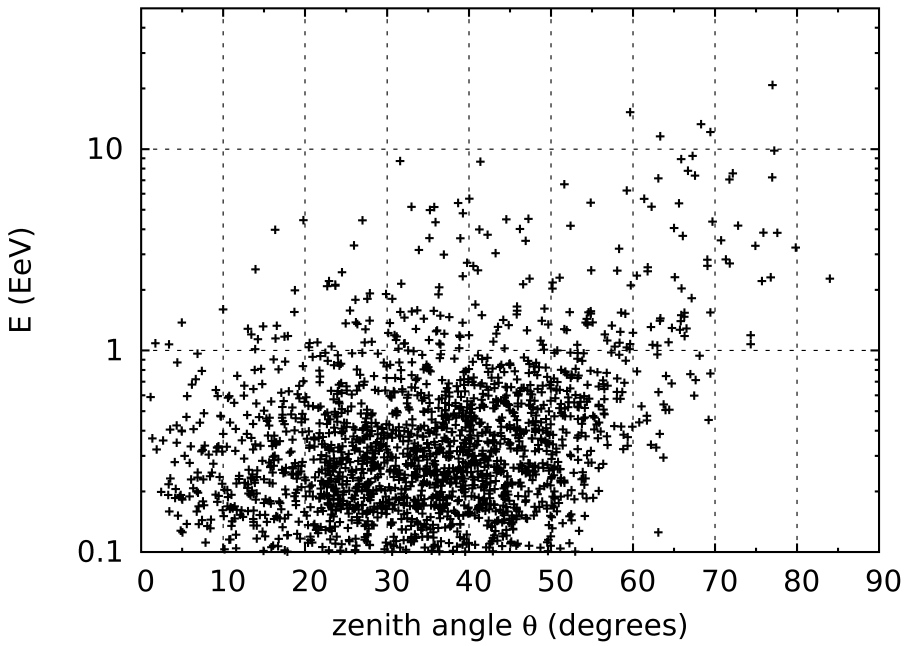

Figure 6.6: Scatter plot of the energy versus the zenith angle of the selected SD events.

We will conclude this section by considering the azimuth angle φ of the selected air showers. A distribution of this variable is plotted in Figure 6.7 (left); a scatter plot versus the energy is shown on the right hand side.

The histogram shows no asymmetry that depends on the azimuth angle; neither does the scatter plot. We do not see the well-known east-west effect of cosmic rays [10, 107]. This effect describes a larger cosmic-ray flux from the west due to the deflection of the predominantly positively charged cosmic rays in the Earth's magnetic field. However it is not surprising that we do not observe this asymmetry as the east-west effect occurs at cosmic-ray energies of around 10 GeV.

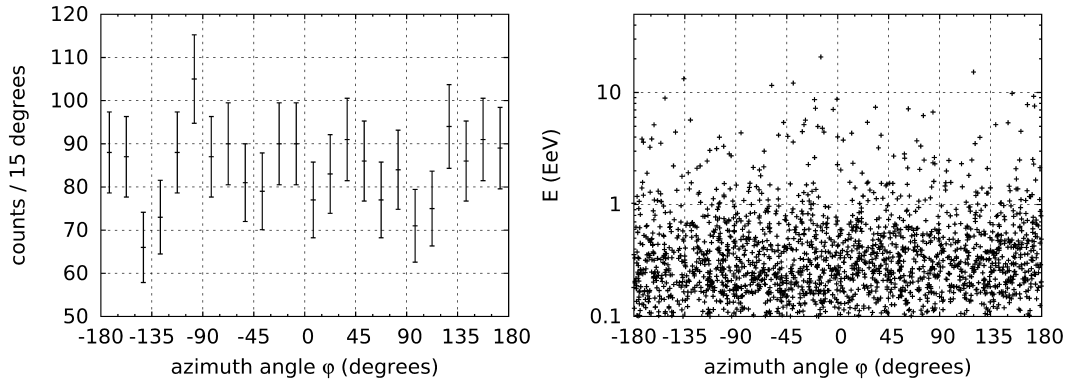

Figure 6.7: Left: distribution of the azimuth angles φ of the selected SD events. The azimuth angle is zero for events coming from the east. Right: scatter plot of the energy versus the azimuth angle.

6.2 Digitizers: features, calibration and timing

With the Nikhef digitizers, two traces with a length of 10 µs each were recorded after each successive trigger. The first trace is channel 1, the measurement of the signal in the north–south direction; the other trace is channel 2 with the data measured in the east–west direction. We use the convention that "north–south direction" means that the antenna arms with which the signal was measured, were pointing to the north and to the south. An example of a trace is shown in Figure 6.11 on p. 81.

The 10 µs trace was sampled with a frequency of 400 MHz; thus each trace consists of 4000 data points. The pre-trigger time was set to 2 µs, which means that the first 800 data points of each trace were measured before the trigger moment. This feature of the digitizer is provided by a digital data buffer.

Inside the digitizer, the input voltage was sampled by a 12-bit Flash ADC (analog-to-digital converter). The measured traces thus consist of 4000 unsigned 12-bit numbers. In order to be able to convert these numbers to voltages, the digitizers were calibrated by applying a 50 MHz 100 mV RMS sine wave to the input terminals. This was done for both inputs of each of the three digitizers individually. Subsequently, sine waves were fitted to the measured traces, from which the calibration parameters were computed. These parameters are listed in Table 6.2.

Digitizer	Channel 1: mV/ADCU	Channel 2: mV/ADCU
1	0.127	0.126
2	0.128	0.128
3	0.126	0.156

Table 6.2: Results of the calibration of the digitizers. Note the different result for digitizer 3, channel 2. ADCU stands for ADC units.

With the exception of the second channel of digitizer 3, the calibration results are very consistent: (0.127 ± 0.001) mV/ADCU. For simplicity, a value of 0.127 mV/ADCU was used in all calculations, with the exception of channel 2 of digitizer 3, where 0.156 mV/ADCU was used. The deviating calibration of this channel is caused by a manufacturing error (a wrong resistor was used).

Not only two traces are stored whenever the digitizer is triggered, but some additional information as well. This information is the event header, and part of it is the GPS timestamp of the event. The timestamp indicates the time at which the digitizer triggered. Due to the pre-trigger time, the time of the beginning of the trace is the timestamp minus the pre-trigger time (typically 2 µs). The first number of the timestamp is the number of seconds in UTC since the UNIX epoch, midnight 1 Jan 1970. The second (and last) number of the timestamp is the number of nanoseconds since the beginning of the second. Logically, this number should range from 0 until 10^9-1.

Uncertainties in the value of the timestamp can be systematic or statistical:

- Systematic uncertainties can arise from the hardware implementation (brand and model of GPS device, cable lengths), incorrect handling of leap seconds, and from counting errors like the bug mentioned in Section 5.3. The advantage of this type of uncertainty is that the errors are constant: we can measure them and compensate for them. The digitizers that were used contain a different GPS timing system than the Auger SD stations, leading to a non-zero offset in timestamp values.
- Statistical uncertainties in the timestamps are different for each event. This type of uncertainty cannot be compensated for, and limits the

precision with which we can determine the exact time of the trigger moment.

To investigate the uncertainty in the timestamp, we analyse the data that was measured within an 8-day period, from midnight 2007-08-08 until midnight 2007-08-16. There is nothing special about this time period; however, it is a convenient period of continuous measurements without any on-site activity.

A search in the data from this period for "events" where traces from all three poles are available, yielded 170,520 such events. Here we used the criterion that the timestamps of the traces from each pole must not differ more than 10 µs from each other. Since a scintillator trigger was used, nearly all events recorded satisfy this criterion. For each of the antennas, 6.3 % of the events could not be fitted into such a threefold coincidence due to the dead time and issues with the data storage. For all other events, i.e. the ones that satisfied the criterion, the time differences of the GPS timestamps from digitizers 1, 2 and 3 have been plotted in Figure 6.8.

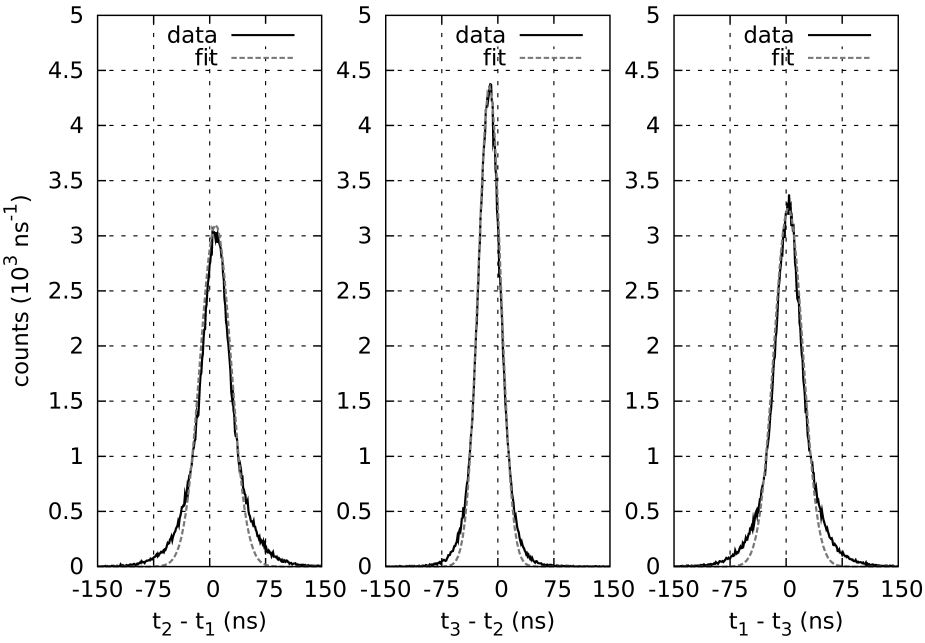

Figure 6.8: Relative time differences of events from the three digitizers, calculated from the GPS timestamps in the event headers. Since the digitizers received the trigger simultaneously, perfect timing would be a time difference of zero.

A Gaussian function of the form $y = A_n \mathrm{e}^{-\frac{(t-t_n)^2}{2\sigma_n^2}}$ may be fitted to each distribution of calculated time differences. We define $n \in \{21, 32, 13\}$, $t_{21} \equiv t_2 - t_1$, and similar for t_{32} and t_{13}. The fit results are summarized in Table 6.3.

Time difference	n	A_n (ns^{-1})	t_n (ns)	σ_n (ns)
$t_2 - t_1$	21	$3.1 \cdot 10^3$	8	21
$t_3 - t_2$	32	$4.4 \cdot 10^3$	−11	15
$t_1 - t_3$	13	$3.3 \cdot 10^3$	3	19

Table 6.3: Fitted parameters for a Gaussian function that was fitted to the distribution of calculated time differences.

Now that the standard deviations of the distributions of the time differences are known, we can calculate the standard deviations of the statistical uncertainties of the GPS timestamps from the individual digitizers. Using $\sigma_{21}{}^2 = \sigma_2{}^2 + \sigma_1{}^2$, etc. and solving the three equations, we find $\sigma_{digitizer}$ = 17, 12, and 9 ns for digitizers 1, 2, and 3 respectively. Furthermore, the average standard deviation of the time difference is $\overline{\sigma_n} = 18\,\mathrm{ns}$, and the average standard deviation of the GPS timestamp of the digitizers is $\overline{\sigma_{digitizer}} = 13\,\mathrm{ns}$. This uncertainty is lower than the uncertainty of 25 ns (1-sigma) that is mentioned in the data sheet of the GPS receiver module [97].

The angular reconstruction of air-shower events is based on the time differences as recorded by the digitizers. The typical uncertainty in the time difference between two GPS timestamps, 18 ns, aggregates into an uncertainty in the angular reconstruction. We can estimate the order of magnitude of this aggregated uncertainty in the angular reconstruction using a simple back-of-the-envelope calculation. Remember that the radio antennas in the setup form an equilateral triangle with a side-length of 100 meters (Section 4.1). Next, consider a fully vertical shower, thus hitting ground level at each of the three antennas at the same time. For simplicity, assume that two of the three digitizers measure this time perfectly. The third digitizer, however, is off by the typical value for the uncertainty in the time *difference*, i.e. 18 ns. In this case, the error in the angle is

simply $\arctan\left(\frac{18\,\mathrm{ns}\cdot c}{½\sqrt{3}\cdot 100\,\mathrm{m}}\right)=4°$. Although the calculation is very simplistic, it gives a quick estimate of the angular uncertainty we can expect based on the limited timing precision of the digitizers.

Updated versions of the digitizers, such as the ones used in the MAXIMA setup, have a higher timing accuracy; see Section 8.2.2.

6.3 Signals in the trace

Most of the traces that have been measured do not show a clear signal; see, for example, the trace in Figure 6.11 on p. 81. Some structure can be seen, however, when the average of a large number of traces is calculated. This is due to the fact that the noise which dominates the trace is random, and therefore will be cancelled when we take the average of a large number of traces. For this analysis, we used data from the same 8-day period as in Section 6.2, i.e. from midnight 2007-08-08 until midnight 2007-08-16. For each of the three antennas and for both the north-south and east-west directions, the average trace was calculated by summing over all traces in the selected period, and then dividing by the number of traces. The result is shown in Figure 6.9.

First it should be noted that only signals that have the same phase-time relationship in each trace, will be visible in this plot (i.e. constructive interference). The plot shows that each of the average traces consists of two or three bursts/peaks. These are not signals from cosmic rays, but noise signals that are created by our own electronics. The first burst, with a peak height of about 1 mV, starts at t=2.0 μs, the trigger moment. Apparently, whenever the digitizer triggers, some noise is generated, which in turn is picked up by the electronics of the setup.

After this initial burst, two additional bursts are visible. These bursts are strong in antenna 1, up to ~10 mV, while weaker but still visible in antennas 2 and 3. For each antenna, they occur at slightly different positions in the trace. However, the time in between these two bursts is always 1.6 μs. This 1.6 μs is the time during which the TTL signal, which is used in the scintillator-based trigger system that is described in Section 4.6, has a value of 'true' after the occurrence of a trigger. Whenever the logical value of the TTL signal changes, there is a

sharp transition in the voltage level. During this transition, radiation leaks out of the coaxial cables, which are located at the data-acquisition (DAQ) system inside the Balloon Launching Station (BLS). This radiation is picked up by the antennas outside. The different times at which the bursts occur in the trace can then be explained by the different distances of the antennas to the BLS, where the DAQ is located. These distances are 30, 117 and 86 m for poles 1, 2 and 3 respectively. Since the noise signal propagates with the speed of light, this takes 0.10, 0.39 and 0.29 µs, respectively. When taking into account the time it takes for the signal to travel back from the antenna over the coaxial cable (835 ns for all antennas) and the pre-trigger time (2.0 µs), we can explain the starting times of these noise bursts. Additionally, the stronger appearance of the noise in antenna 1 can simply be explained by the fact that this antenna is nearest to the BLS.

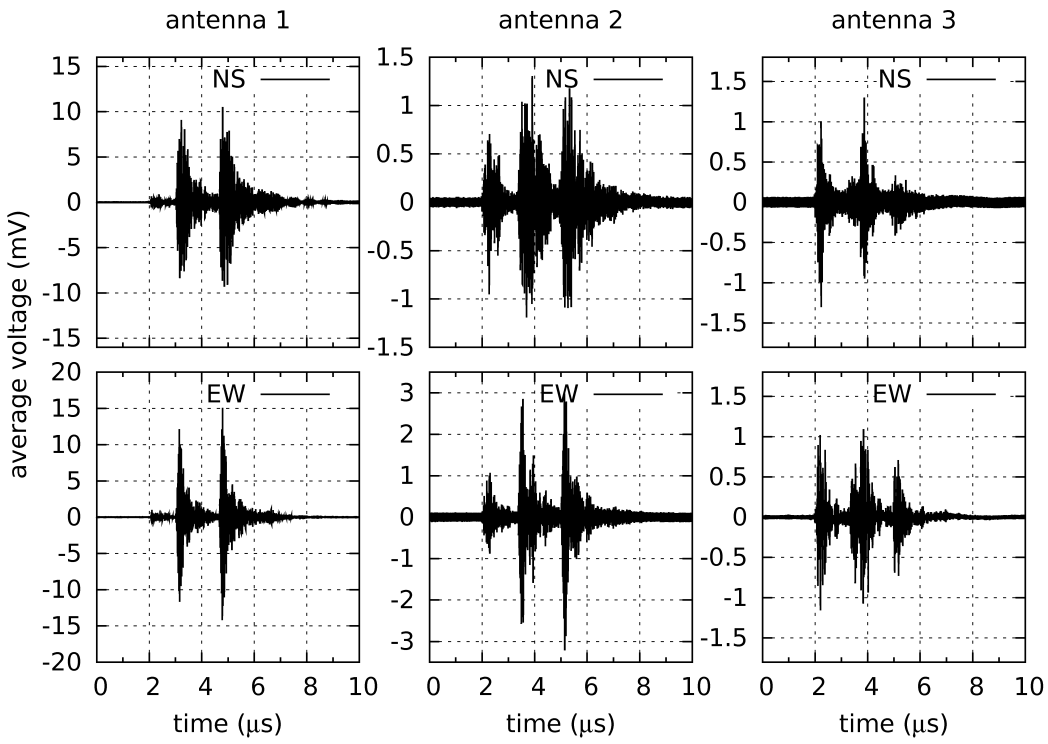

Figure 6.9: Average of about 182,000 traces for each of the three antennas, north-south and east-west directions, calculated for a period of 8 days.

Now that the timing of the noise in the trace due to self-interference is understood, it can be compared with the timing of the signal. Where in the trace do we expect to see radio signals from air showers? Let us define t_0 as the time at

which the muons in the air shower pass through the scintillators located at the BLS. After that, it takes the scintillator signal 100 ns to reach the trigger unit, the trigger unit itself takes 50 ns, and finally the trigger distributor ('fan-out' device) introduces a delay of 12 ns. Therefore, at t_0 + 162 ns the trigger reaches the trigger input of the digitizers.

The signal itself traverses the cable between the antenna and the BLS in 835 ns; this is true for all three antennas. However, since the antennas are not located at the BLS, the shower front may not pass the antennas at t_0. When ignoring the curvature of the shower front, the front passes the antennas at t_0 if the event is fully vertical; i.e. the zenith angle is zero. When the zenith angle is not zero, then the arrival time depends on the zenith and azimuth angles of the shower and on the distance of the antenna to the scintillators. These expected arrival times are summarized in Table 6.4 for the three antennas and three possible zenith angles.

antenna	zenith angle θ	shower front at antenna (ns)		signal at digitizers (ns)		signal in the trace (ns)	
		earliest	latest	earliest	latest	earliest	latest
1	0°	t_0	t_0	t_0+835	t_0+835	2,673	2,673
2	0°	t_0	t_0	t_0+835	t_0+835	2,673	2,673
3	0°	t_0	t_0	t_0+835	t_0+835	2,673	2,673
1	60°	t_0-87	t_0+87	t_0+748	t_0+922	2,586	2,760
2	60°	t_0-338	t_0+338	t_0+497	t_0+1,173	2,335	3,011
3	60°	t_0-250	t_0+250	t_0+585	t_0+1,085	2,423	2,923
1	90°	t_0-101	t_0+101	t_0+734	t_0+936	2,572	2,774
2	90°	t_0-390	t_0+390	t_0+445	t_0+1,225	2,283	3,063
3	90°	t_0-288	t_0+288	t_0+547	t_0+1,123	2,385	2,961

Table 6.4: Times at which the signal of the shower can be expected, calculated for three antennas and three possible zenith angles. The earliest and latest possible times are given; the actual times depend on the geometry of the shower. In the calculation, a flat plane approximation was used for the shower front.

Instead of calculating the average voltage for each point in the trace, one may also calculate the average power for each point. Since $P = V^2/R$, this amounts to summing the squares of the voltages, then dividing by the number of traces and

R. The difference with the average traces as shown in Figure 6.9 is that signals with a random phase-time relationship are no longer cancelled. The calculated "average powers" are shown in Figure 6.10.

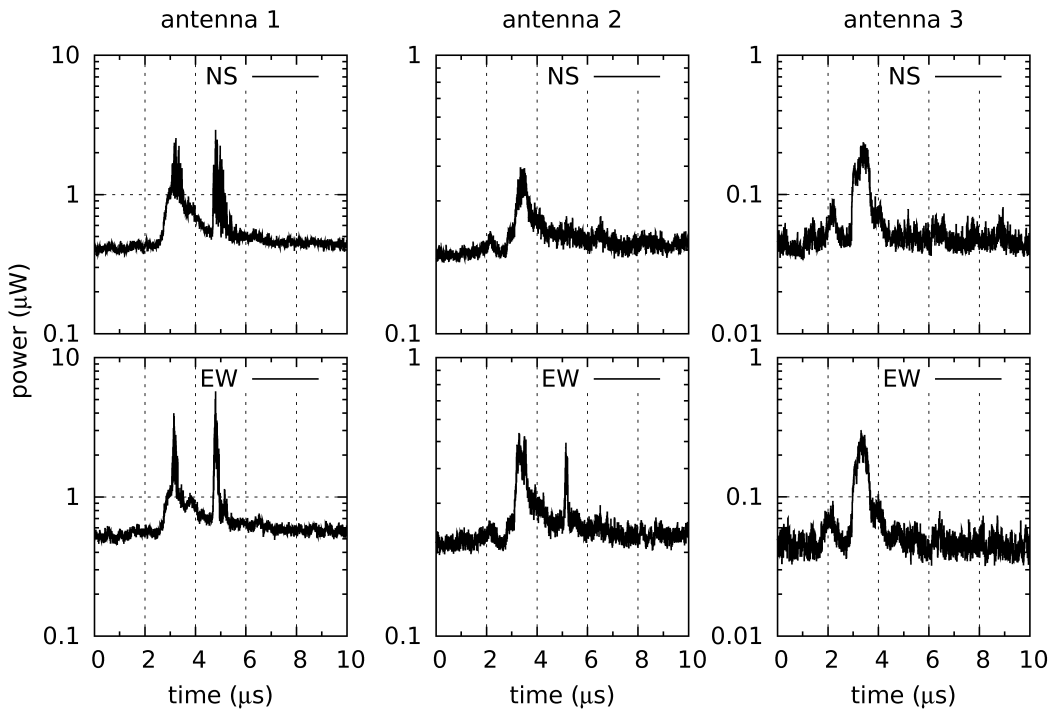

Figure 6.10: Average power for each point in the trace. Calculated for each of the three antennas, for the north-south and east-west directions, for a period of 8 days.

When comparing Figure 6.9 with Figure 6.10, we see that both figures show compatible results for antenna 1. In Figure 6.9, the two bursts in the signal from antenna 1 are well above 5 mV, and in Figure 6.10 two corresponding peaks are visible. This is the self-interference from the TTL trigger signal. For antenna 2 however, Figure 6.10 shows two peaks in the east-west direction, but only one peak in the north-south direction. When looking in Figure 6.9 at the plot for antenna 2, north-south, we see that the TTL-induced burst is not very strong; around 1 mV. The same is true for antenna 3 in both the north-south and east-west directions. This may explain why no TTL-induced bursts are visible above the noise in the corresponding plots in Figure 6.10. As was explained above, noise with a random phase is cancelled in Figure 6.9, but not in Figure 6.10.

6.4 Properties of the radio background

To discuss the radio background, we will investigate the noise as measured by the first-generation radio setup. The noise is the background upon which we may measure signals.

Within the measured traces, at least three kinds of noise can be distinguished. First there is the general background, present at all times and at all frequencies. However, it is not equally strong at all frequencies, nor is it constant over time. Two examples of this type of noise are thermal noise and galactic noise. Second is the narrow-band noise, which is a continuous signal transmitted at a single frequency. Naturally, there can be many sources of narrow-band noise active at one moment, resulting in many peaks in the frequency spectrum. An example of this is the various frequencies in the FM broadcast band. The third kind of noise we can distinguish are the so-called transients, which are short bursts of broad-band noise. This type of noise will be discussed in more detail in Section 6.4.1.

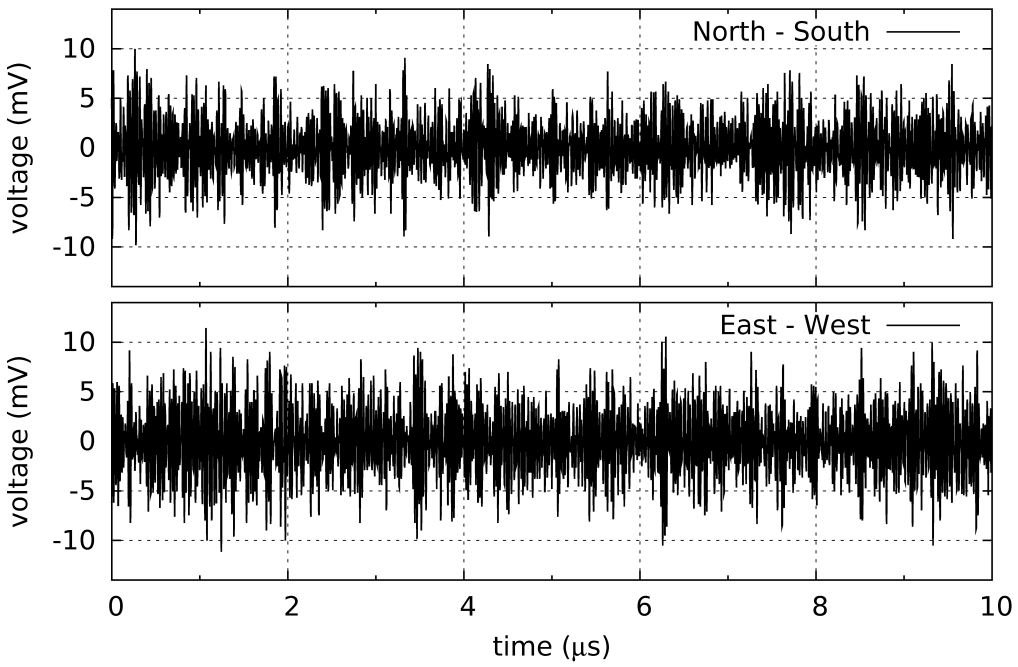

Figure 6.11: Example of a trace containing noise. Traces have a length of 10 µs and consist of 4000 points. The trace shown in this figure was measured on August 8, 2007.

To begin the discussion of noise, an example of a trace with noise has been plotted in Figure 6.11. Very few features can be seen in the example trace; the measured signal is continuous and fairly random. More details can be seen in the frequency spectrum of the trace, which was calculated using a 4096-point fast Fourier transform; see Figure 6.12.

The sampling frequency of the digitizer was 400 MHz; therefore the Nyquist frequency of the recorded data is 200 MHz. What has been plotted in Figure 6.12 is the first Nyquist band, 0–200 MHz. On the vertical axis is plotted the root mean square (RMS) voltage for each frequency, which is $\sqrt{2}$ times smaller than the amplitude. However, it is more common to express the frequency spectrum in terms of power in decibels (dBm). The logarithmic "dBm" unit expresses the power of a signal relative to a reference power of 1 mW. (For example, a signal with a power of 10^3 mW is a signal of 3·10 dBm.)

The frequency spectrum can also be represented using the power spectral density (PSD). The PSD is closely related to the power, and expresses the amount of power per unit of frequency.

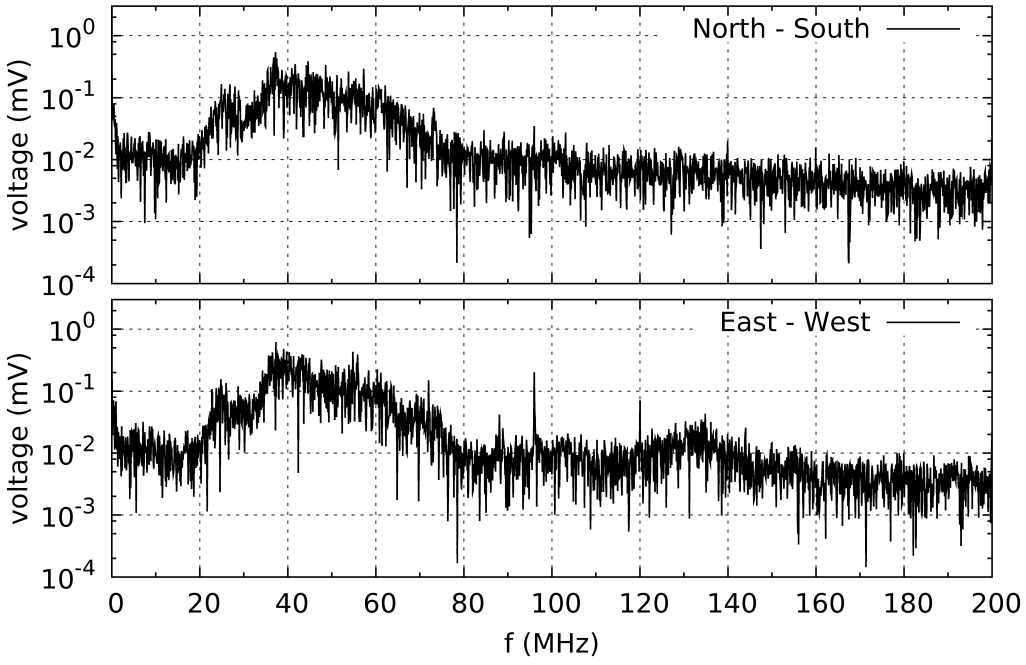

Figure 6.12: Frequency spectrum of the trace shown in Figure 6.11.

For narrow-band signals, like FM broadcasts on a “single” frequency, it is logical to work in units of power, as one may talk of “the power of a signal”. Broad-band signals, on the other hand, are naturally expressed in terms of power spectral density. The power of these signals is spread across many frequencies. What can be measured in real life, however, is the amount of power within a frequency bin. When using small bins, the amount of power within a bin becomes proportional to the size of the bins. Therefore, the power of a broad-band signal at a certain frequency is not defined. (It would depend on the size of the bins of the measurement device that one is using.) However, what can be defined is the density of the power, which can be calculated by taking the power within a bin, and dividing by the size of that bin. Therefore, the PSD is expressed in units of dBm/Hz or dBm/MHz. For broad-band signals, the PSD does not depend on the spectral resolution (the size of the frequency bins). This allows the comparison of measurement results. For narrow-band signals, on the other hand, the PSD resembles the shape of a Dirac delta function and is not very useful to describe the signal.

Figure 6.12 shows that nearly all of the power in the trace is contained within a frequency band of roughly 20 – 80 MHz. In the east-west polarisation, a source of narrow-band radio-frequency interference (RFI) is present at 96 MHz, which lies within the FM broadcast band.

Although some structure can be seen in Figure 6.12, many details are hidden underneath the random noise. To overcome this, we have averaged the power spectra of all traces ($1.8 \cdot 10^5$) measured within an 8-day period. Again, we use the same 8-day period as in Section 6.2. The 8-day averaged power spectrum is presented in Figure 6.13.

In comparison with the previous figure, many more details can be seen in this plot. The RFI peak at 96 MHz is now very visible, as well as additional peaks at 21.7, 48.2, 55.3, 72.0, and 120.0 MHz, amongst others. Also visible is the effect of the high-pass and low-pass filters. The attenuation of the filters is a function of frequency and changes gradually: from -3 dB at the cutoff frequencies f_{co} (25 MHz and 67 MHz, respectively) until -20 dB in the stop band (< 19 MHz and > 90 MHz). See Section 4.3.

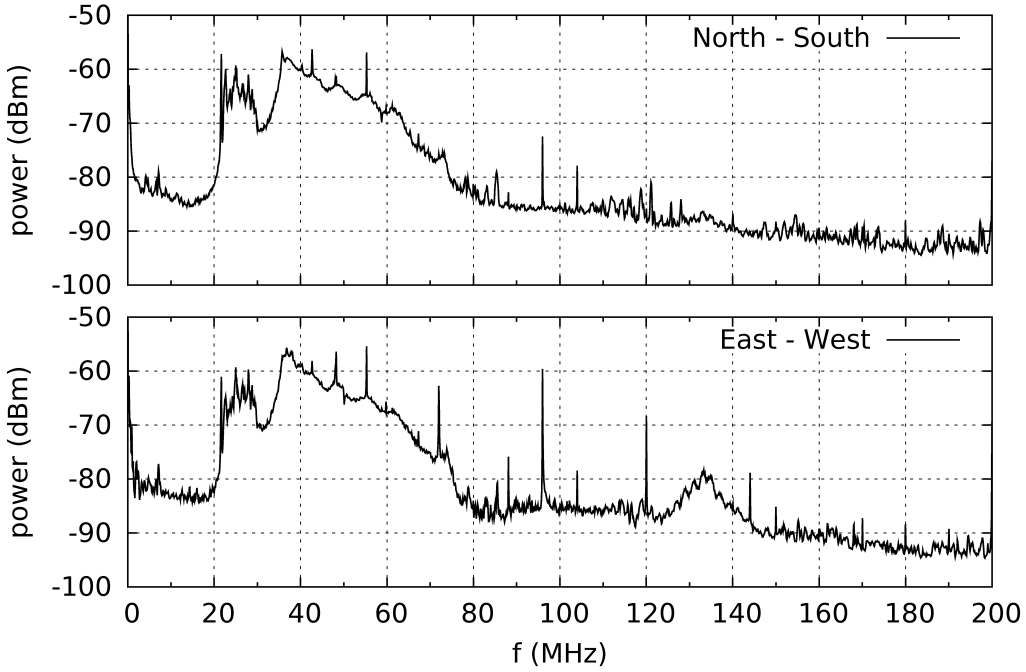

Figure 6.13: Average power spectrum of all traces measured during 8 days.

Another feature that is visible in the power spectrum is the antenna response. The spectrum in Figure 6.13 falls off below 36 MHz. This frequency is comparable with the resonance frequency of the longest arm of the antenna, 40 MHz. Figure 4.5 on p. 45 shows that the power transfer factor of the aluminium LPDA is 0.5 (i.e. −3 dB) at 36 MHz; this is in agreement with the power spectrum in Figure 6.13. Additionally, the resonance frequencies of the individual arms of the antenna are somewhat visible.

Another detail in Figure 6.13 is the small bump that can be seen around 130 MHz in the east–west direction. This is not actually a signal picked up by the antenna, but merely an artefact of the digitizer.

As discussed in Section 6.3, one aspect of the radio setup is the pollution of the measured traces with noise caused by the TTL-trigger signal. Additionally, since the setup was designed to measure radio signals from air showers, some of these signals may be present in the traces. For these reasons, the selected traces are not a 100 % unbiased sample of the radio background.

To eliminate these artefacts, we can choose not to use the full 10 µs of the

trace, but only the pre-trigger part. The pre-trigger part is the part of the trace that has been measured before the trigger moment, which in this case is the first 2 µs of the trace. Considering the sampling frequency of 400 MHz, this amounts to the first 800 samples of each trace. Since fast Fourier transforms work best on arrays with a length that can be written as a power of 2, we will use the first 512 samples of each trace.

The result of the recalculation of the average power spectrum, now using only the first 512 points of the trace, is presented in Figure 6.14. For comparison, the original results have been added in grey. Since the two results have a different spectral resolution, we are comparing the power spectral density instead of the power itself.

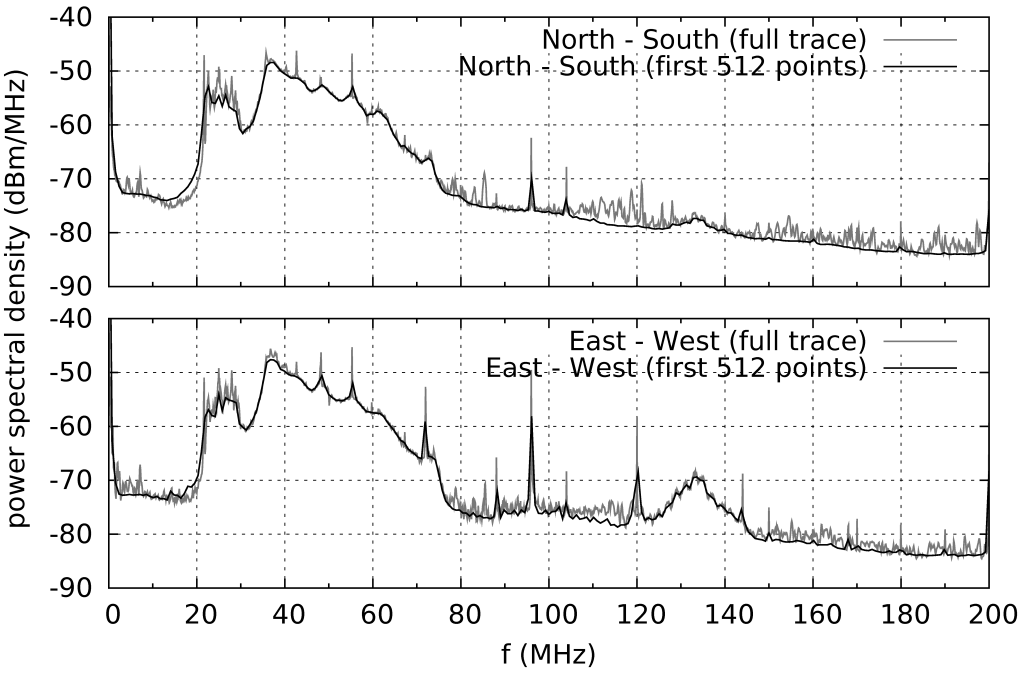

Figure 6.14: Average power spectral density of the traces measured during the 8 day period. The biased data of the full trace (grey) can be compared with the unbiased data of the partial trace (black). Since the partial trace has a lower resolution, peaks in the spectrum are less clearly defined.

When comparing the two spectra in Figure 6.14, few differences can be seen. Narrow-band RFI peaks are less sharp in the spectrum from the partial trace, which is to be expected due to its lower spectral resolution. At frequencies at which there is a sharp drop-off in the spectrum, e.g. at 20 MHz, the spectrum

from the partial trace changes its value more slowly than the full spectrum. In at least two frequency bands, 100–120 MHz and 145–200 MHz, the spectrum from the full trace is more noisy than that of the partial trace. Although there is a clear difference between both spectra at these frequencies, the power level is below −70 dBm/MHz, thus relatively low.

The spectra in Figure 6.14 were calculated by averaging over a relatively long period of 8 days. It is also possible to calculate an average spectrum over a shorter period, e.g. 5 minutes, and then consider how the spectrum changes as a function of time. This is the procedure that was used to calculate the spectra in Figures 6.15 and 6.16, using the said period of 5 minutes. In this calculation, the full traces were used, as opposed to using only the first 512 points of each trace. The reason for this is the superior spectral resolution of spectra calculated from full-length traces. From Figure 6.14 it can be concluded that the noise in the spectra due to self-interference is not very large.

Figures 6.15 and 6.16 show that the intensity of the spectrum between 25 and 75 MHz changes with time, with a period of approximately 24 hours. It was shown by Coppens [108] that the period of this cycle is not exactly 24 hours, but one sidereal day (23.93 hours). A sidereal day is the period required for the Earth to complete one rotation around its axis relative to a fixed background of stars. The sidereal day is slightly shorter than a regular day due to the rotation of the Earth around the Sun. (In fact, $24 \cdot \frac{364}{365}$ hours is a good approximation for the length of a sidereal day.

The important conclusion of the periodicity of the spectrum with the sidereal day is that a large part of the signals in the spectrum are of galactic origin. The antenna setup is sensitive enough to measure the galactic background above the noise. This noise, which includes self-made noise, man-made noise, atmospheric noise, and other sources of noise, is thus relatively low.

Sources of RFI that were visible in the spectra of Figure 6.13, such as the peak at 96 MHz, can be seen as vertical lines in Figures 6.15 and 6.16. Again, this signal is stronger in the east-west direction than in the north-south direction. Horizontal lines seen in Figures 6.15 and 6.16 represent periods during which the noise levels were temporarily higher. Since these lines span a broad frequency range, they represent broad-band noise. Possible candidates for this

type of noise are man-made noise and noise generated by thunderstorms, amongst others.

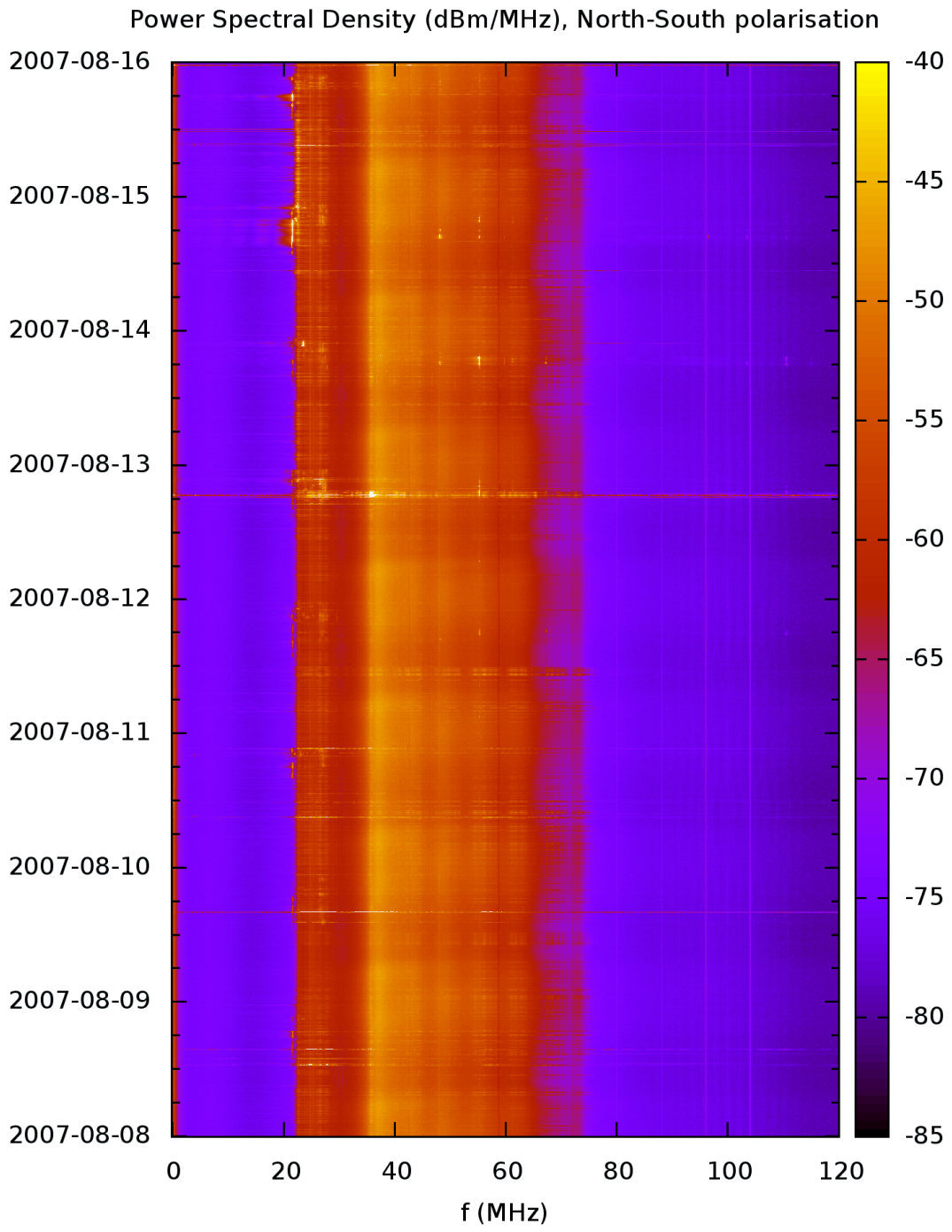

Figure 6.15: Power spectral density (PSD) in the north-south polarisation, as measured during 8 days in August 2007.

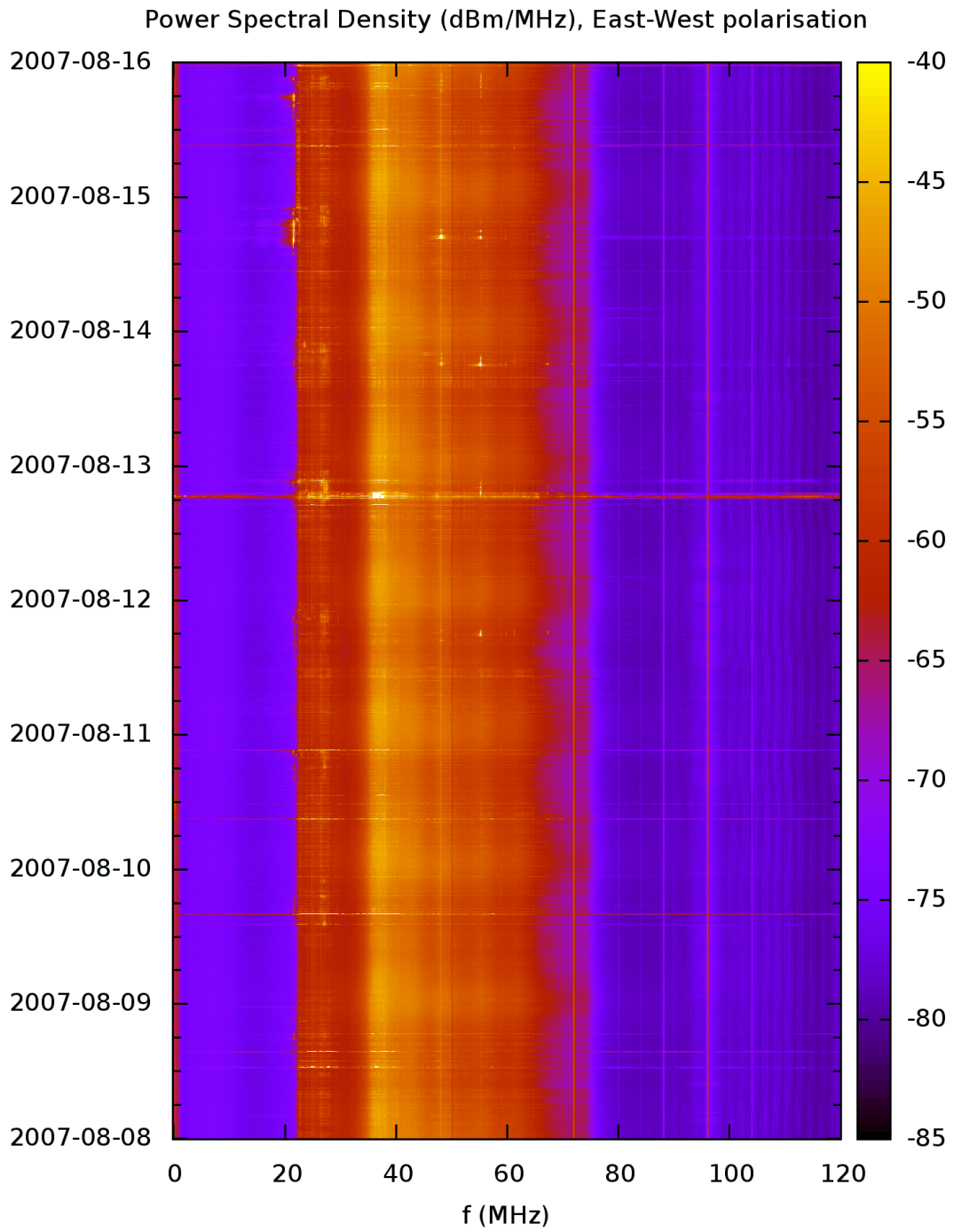

Figure 6.16: Power spectral density (PSD) in the east–west polarisation, as measured during 8 days in August 2007.

6.4.1 Transients: short bursts of noise

The horizontal lines in Figures 6.15 and 6.16 were discussed as periods during which a large amount of broad-band noise was received. However, broad-band noise is also received outside of those periods. Short bursts of noise, the so-called transients, are received several times per second. There is a large variation in the amplitude and shape of the transients, as well as the frequency with which they are received. Depending on the environmental conditions and the chosen threshold value, this repetition frequency of transients can be very high.

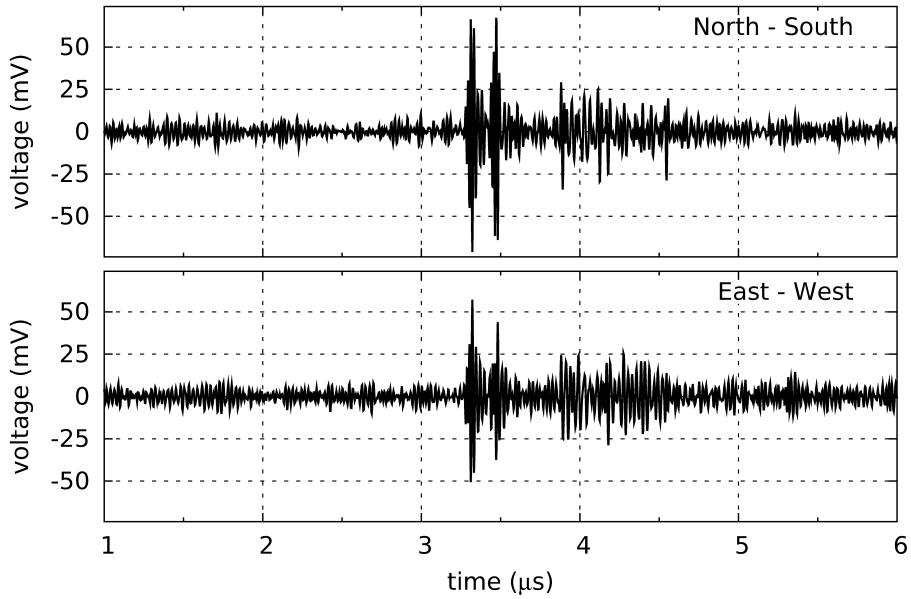

Figure 6.17: Example of wide-band noise in a trace

Transients can have various shapes; from long bursts of noise, sometimes repetitive, to short bursts, similar to the example transient in Figure 6.17. Three additional examples of transients are shown in Figure 6.18; in Figure 6.18 (b) saturation of the trace can be seen. During saturation, the amplitude of the signal is larger than the largest voltage that the ADC can digitize, or smaller (more negative) than the smallest voltage that the ADC can digitize.

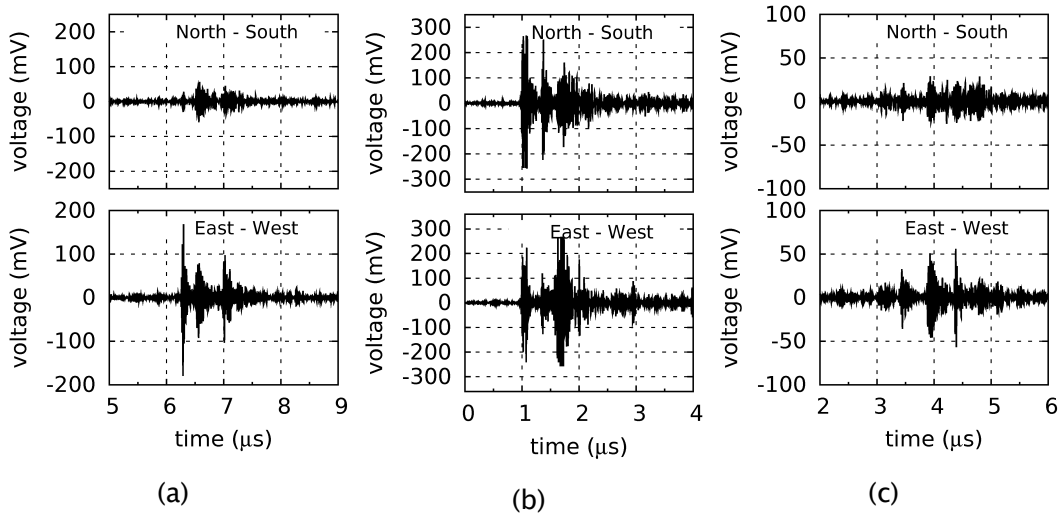

Figure 6.18: More examples of transients.

6.5 Distribution of amplitudes in the traces

Another method to analyse the radio background is to look at the distribution of the voltage values in the measured traces. Each trace consists of 4,000 data points (ADC values) that represent the input voltage of the digitizer, with samples taken every 2.5 ns. Using the calibration result of (0.127 ± 0.001) mV/ADCU (Section 6.2), these ADC values can be converted to voltages.

Because the DAQ system contains a bias-T and high-pass filters, no DC signal is received from the antennas. Since all ADC values are positive, the 0 mV point is determined by subtracting the mean value of the trace. This is done for each trace individually.

For each trace, one can make a histogram of the voltage values of the 4,000 data points in the trace. Such a histogram shows the distribution of the voltage values within the trace, i.e. the frequency with which each voltage value occurs. This distribution is called the amplitude probability distribution (APD).

Instead of creating a histogram using the data points from a single trace, one can obtain better statistics by using a large number of traces. This has been done for the APD in Figure 6.19, where we have used all the traces that were measured

within the 8-day period that was used previously. The histogram was created with a bin size of 1 ADC unit; this equals 0.127 mV.

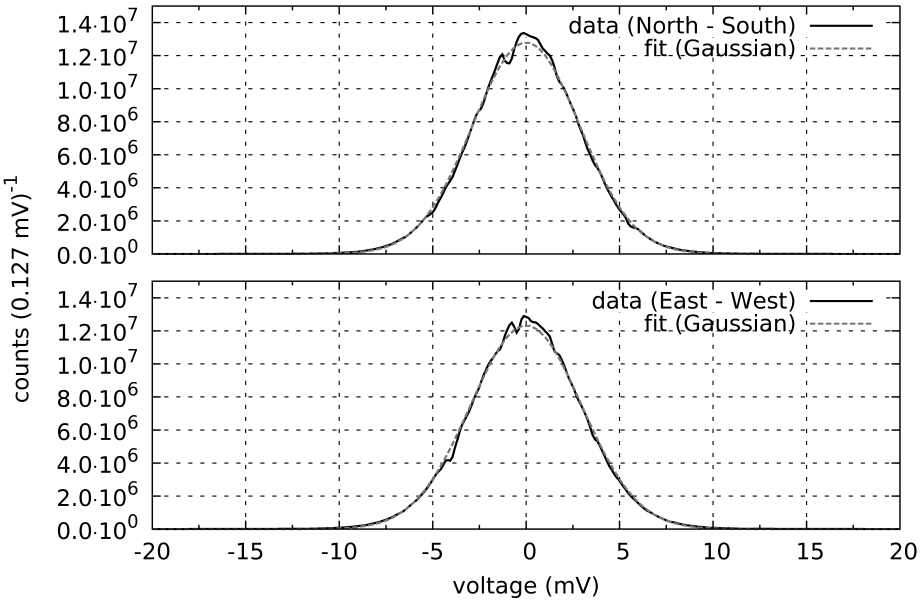

Figure 6.19: Non-normalized amplitude probability distribution (APD) on a linear scale. The histogram was compiled from all traces that were measured during 8 days in August 2007. A Gaussian function has been fitted to the data.

The APD in Figure 6.19 resembles a Gaussian distribution, and a Gaussian function $y = A\,\mathrm{e}^{-\frac{(V-V_0)^2}{2\sigma^2}}$ has been fitted to the distribution. Because of its definition, we have forced $V_0 = 0$. Since the function fits rather well, a large part of the measured noise can be considered Gaussian noise.

However, the tails of the distribution at $|V| \geq 10\,\mathrm{mV}$ are not described well by the Gaussian function; see Figure 6.20. A different function is required to adequately describe the tail of the distribution; in this case we have pragmatically chosen a modified Breit-Wigner function. The second function that was fitted is $y = A_1\,\mathrm{e}^{-\frac{(V-V_0)^2}{2\sigma^2}} + \frac{A_2}{\pi}\frac{\gamma^n}{|V-V_0|^n+\gamma^n}$. However, other functions may also be chosen to fit the non-Gaussian part of the noise; see e.g. [109].

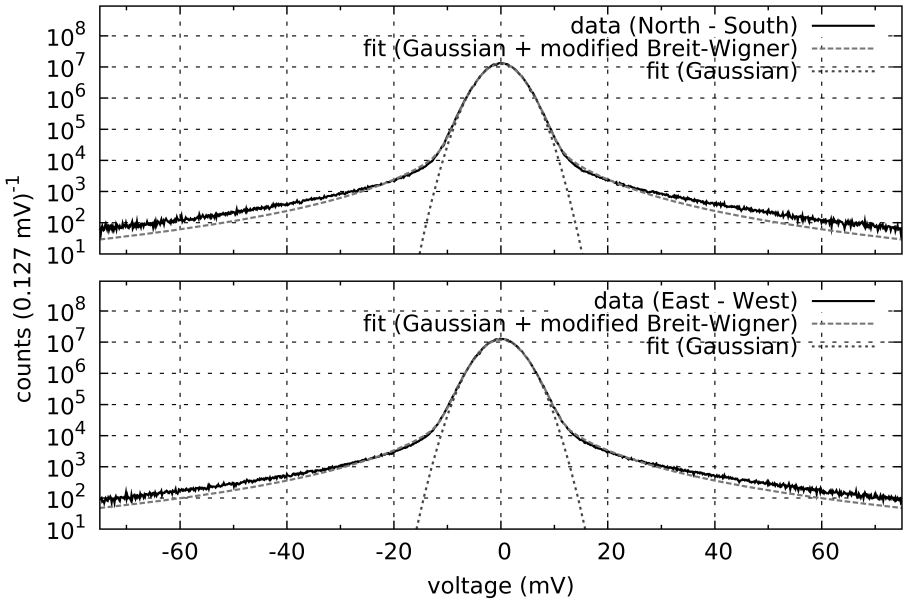

Figure 6.20: The same APD as in Figure 6.19, but then on a logarithmic scale. Two functions have been fitted: a regular Gaussian, and a Gaussian added with with a modified Breit–Wigner function.

Thus, the measured data contains not only Gaussian noise, but also noise of a different kind. Possible forms of this non-Gaussian noise are transients or noise that is generated by the DAQ system itself (i.e. self-interference).

A known source of noise that is generated by the DAQ system is the TTL signal of the trigger; this was discussed in Section 6.3. As a check, one can compare the APD of Figure 6.20 with an APD that has been created from the first 512 data points (~1.3 μs) of each trace. This part of the trace has been measured before the trigger moment (see Section 6.4). Therefore it does not contain noise that was generated after the trigger moment. (However, it may still contain other forms of self-generated noise.)

A comparison of the full APD with the “first 512 points” APD is shown in Figure 6.21. To aid the comparison, all curves have been smoothed with a spline function, as well as normalized to their amplitude at V=0. The “first 512 points” APD has a tail that is slightly below that of the full APD. This indicates that there is less noise (or signal!) before the trigger moment. However, most of the noise is not caused by the trigger. This conclusion is in agreement with the discussion in Section 6.4, where the power spectral densities were compared.

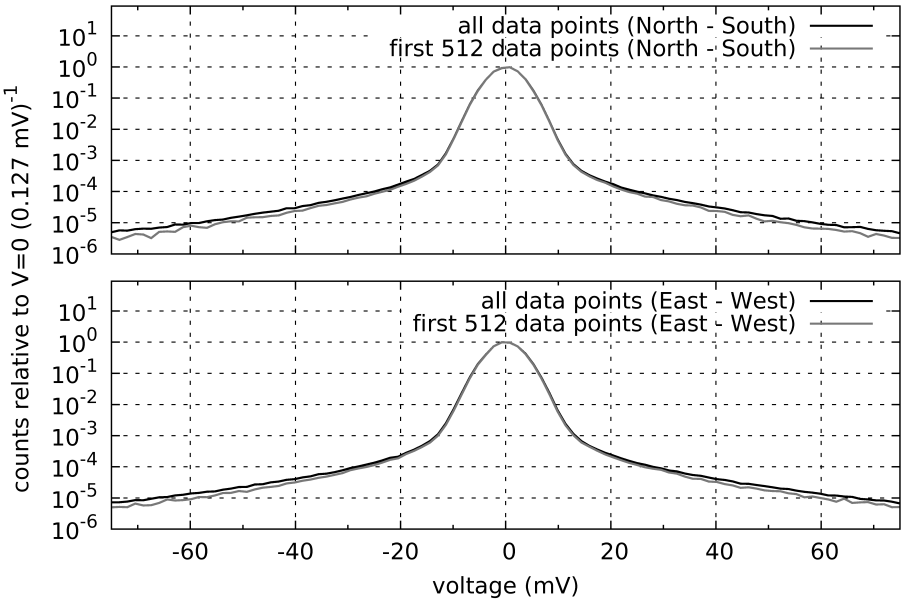

Figure 6.21: Comparison of the APD that was generated from the full trace with an APD that only uses pre-trigger data.